테일러 급수와 뉴턴법

1. 테일러 급수의 이해

테일러 급수 또는 테일러 전개는 어떤 미지의 함수 $f(x)$를 아래와 같이 근사 다항 함수로 표현하는 것을 의미한다.

$$f(x) = p_\infty(x)$$

$$\begin{aligned} p_n(x) &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \end{aligned}$$

테일러 급수에서 주의해야 하는 사항은 좌변과 우변이 모든 x에 대해서 같은 게 아니라 $x=a$ 근처에서만 성립한다는 점이다. 즉, x가 a로부터 멀어질수록 $f(x)=p(x)$로 두는 것은 큰 오차를 갖게 된다.

테일러 급수는 결국 $x=a$에서 $f(x)$와 동일한 미분계수를 갖는 어떤 다항함수로 $f(x)$를 근사시키는 행위이다. 테일러 급수를 이용해 이와같이 $x=a$에서 미분계수를 일치시키면 $x=a$ 뿐만 아니라 그 주변의 일정 구간에서도 $f(x)=p(x)$가 성립하게 된다.

2. 테일러 정리

테일러 정리는 n번 미분 가능한 함수 $f(x)$에 대해서

$$f(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + h_n(x)(x-a)^n$$

$$\lim_{x \to a} h_n(x) = 0$$

을 만족하는 실함수 $h(x)$가 반드시 존재한다는 정리이다.

즉, 테일러 정리는 어떤 함수를 유한한 차수(n차)의 다항함수로 근사할 수 있는 수학적 근거를 제시하고, 이 때 $h(x)(x-a)^n$은 근사오차를 나타낸다.

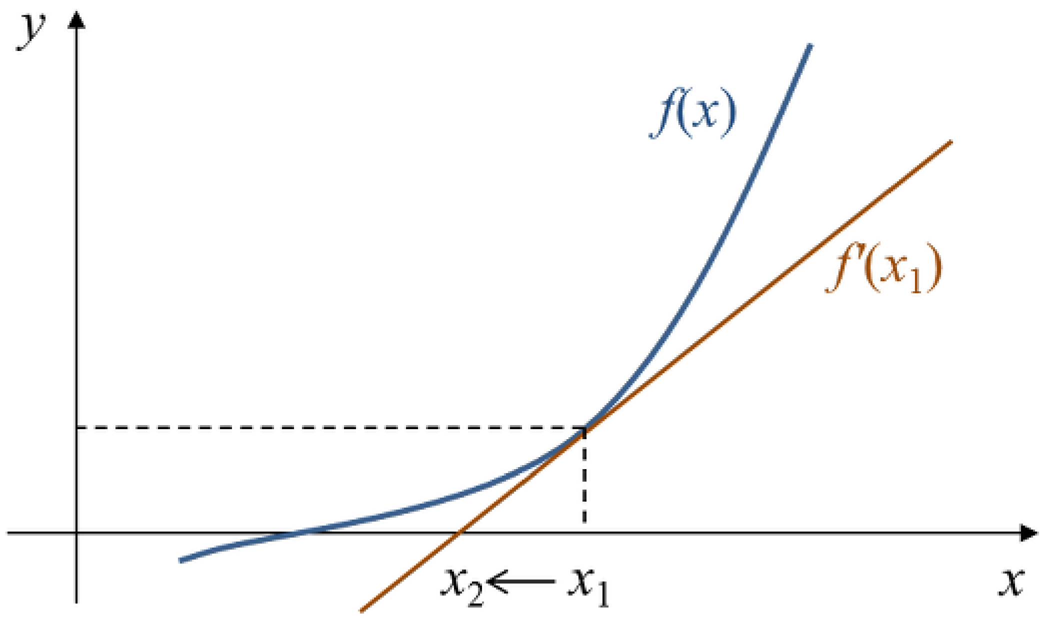
3. 뉴턴법의 이해

뉴턴법은 뉴턴-랩슨법이라고도 불리며, 방정식 $f(x) = 0$의 해를 근사적으로 찾아내는 데 쓰이는 방법이다.

쉽게 위 방정식의 해를 찾기 어려운 함수가 주어졌을 때 뉴턴법은 $f'(a)$가 $x = a$에서 접선의 기울기라는 미분의 기하학적인 해석을 활용한다.

일단 무작위의 $x = a$를 상정하고 $f(a)$의 값을 본다. 만약 $f(a) > 0$이고 이 때 $f'(a) > 0$이라면 실제 값인 x는 a보다 작은 값일 것이다.

이러한 과정을 반복하며 뉴턴법은 현재 x에서 접선을 그리고 그 접선과 x축이 만나는 지점으로 x값을 업데이트 하며 점진적으로 해를 찾아 나간다. 아래 그림을 보면 쉽게 이해가 될 것이다.



이 과정을 수식적으로 표현하면 아래와 같다.

$$x^{t+1} = x^t - \frac{f(x^t)}{f'(x^t)}$$

이 때 업데이트의 종료 조건은 x값의 변화가 거의 없을 때, 즉 $|x^{t+1} - x^t|$가 매우 작은 값이 될 때 종료하고 이 때 $f(x^{t+1}) = 0$이라고 생각하는 것이다.

4. 가우스-뉴턴 방법을 이용한 연립방정식 해결

아래와 같이 변수의 개수가 n이고 식의 개수가 m $(m \geq n)$인 다변수 연립방정식이 존재한다고 해보자.

$$\begin{gathered} f_1(X_1, X_2, \cdots, X_n) = 0 \\ f_2(X_1, X_2, \cdots, X_n) = 0 \\ \cdots \\ f_m(X_1, X_2, \cdots, X_n) = 0 \end{gathered}$$

이를 행렬꼴로 바꾸기 위해 $\boldsymbol{X} = (X_1, \cdots, X_n)$, $F(\boldsymbol{X}) = (f_1(\boldsymbol{X}), \cdots, f_m(\boldsymbol{X}))$로 놓으면 위의 다변수 연립방정식은 $F(\boldsymbol{X}) = 0$꼴이 된다.

이 때 F는 $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$인 다변수 다차원의 함수이며, 따라서 이에 대한 미분은 Jacobian으로 표현된다.

$$J(\boldsymbol{X}) = \begin{bmatrix} \frac{\partial f_1}{\partial X_1} & \cdots & \frac{\partial f_1}{\partial X_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial X_1} & \cdots & \frac{\partial f_m}{\partial X_n} \end{bmatrix}$$

이제 F를 함수, $\boldsymbol{X}$를 변수로 보고 뉴턴법 식을 그대로 쓰면 아래와 같다.

$$\boldsymbol{X}^{k+1} = \boldsymbol{X}^k - \frac{F(\boldsymbol{X}^k)}{J(\boldsymbol{X}^k)}$$

그러나 행렬로 직접 나누는 것은 불가능하기 때문에 대신에 $JP = F$를 만족하는 P를 구함으로써 계산해준다. 결국 연립방정식에 대한 뉴턴법 수식은

$$\begin{gathered} J(\boldsymbol{X}^k)P = F(\boldsymbol{X}^k) \\ \boldsymbol{X}^{k+1} = \boldsymbol{X}^k - P \end{gathered}$$

즉, 현재 우리가 추정한 임의의 $\boldsymbol{X}$에 대한 추정값에 대해 J와 F를 계산하고, $JP = F$인 P를 이용해서 $\boldsymbol{X}$가 수렴할 때까지 업데이트한다.

가우스-뉴턴 방법의 매 단계에서 $F(\boldsymbol{X})$를 테일러 근사하면 $F(\boldsymbol{X}) \approx F(\boldsymbol{X}^k) + J(\boldsymbol{X}^k)(\boldsymbol{X} - \boldsymbol{X}^k)$가 되고, 결국 $F(\boldsymbol{X}) = 0$을 만족하는 $\boldsymbol{X}$는 $F(\boldsymbol{X}^k) + J(\boldsymbol{X}^k)(\boldsymbol{X} - \boldsymbol{X}^k) = 0$을 만족하는 $\boldsymbol{X}$이어야 한다.

따라서 위 식을 $\boldsymbol{X}$에 대해서 풀면

- J의 역행렬이 존재하는 경우: $\boldsymbol{X} = \boldsymbol{X}^k - J^{-1}F$
- J의 역행렬이 존재하지 않는 경우: $\boldsymbol{X} = \boldsymbol{X}^k - (J^\top J)^{-1}J^\top F$

임을 알 수 있다.